

### 1. 指定值

用 `=` 或是 `<-` 即可指定值，例如 `x=4`，`x<-4`，而 `y<-6:10` 即 `y` 值有 6,7,8,9,10，而 `z<- " apple "` 即 `z` 代表字串 `apple`。  
`print(x)`，或是直接輸入 `x`，可以列印值。

### 2. 確認 R 記憶體/刪除 R 記憶體

`ls()`

`rm(x)`

### 3. 四則運算與平方與根號

`sqrt(x)` ( 或是 `x^(1/2)` )：對 `x` 取根號

`exp(x)`：`e^x`

`log2(x)`：對 `x` 取底數為 2 的 `log`

`abs(x)`：取絕對值

\*若算是輸入不全，會出現 `+`，直到完成算式才會消失。

### 4. 註解

`#`

### 5. 向量、兩字串、連續整數、等差數列、重複

`c(1,3,5,7)`：向量 1,3,5,7

`c( " male " , " female " )`：字串 `" male "`、`" female "`

`2:7`：連續整數 2,3,4,5,6,7

`seq(from=1, to=7, by=1/2)`：等差數列，首項為 1，末項為 7，公差為 1/2。

`rep(1:3, times=10)`：1 2 3 重複 10 次

### 6. 數列四則運算

ex: `x=1:5`

`x+10`

`>11 12 13 14 15`

## 7. 向量的四則運算

ex: `x=1:5`

`y=c(1,3,5,7,9)`

`x+y`

`>2 5 8 11 14`

\*當兩者長度相同才可以這樣使用。

## 8. 從向量中選取數值

ex: `y=c(1,3,5,7,9)`

`>1 3 5 7 9`

取出第四項數值

`y[4]`

`>7`

取出所有 ( 除了第四項 ) 數值

`y[-4]`

`>1,3,5,9`

取出第一項和第五項數值

`y[c(1,5)]`

`>1 9`

取出 `y<6` 的數值

`y[y<6]`

`>1 3 5`

## 9. 矩陣

`matrix(c(1,2,3,4,5,6,7,8,9), nrow=3, byrow=TRUE)`

```
> matrix(c(1,2,3,4,5,6,7,8,9), nrow=3, byrow=TRUE)
     [,1] [,2] [,3]
[1,]    1    2    3
[2,]    4    5    6
[3,]    7    8    9
```

3\*3 矩陣 · 按橫排排列

```
> matrix(c(1,2,3,4,5,6,7,8,9), nrow=3, byrow=FALSE)
      [,1] [,2] [,3]
[1,]    1    4    7
[2,]    2    5    8
[3,]    3    6    9
```

3\*3 矩陣，按直行排列

#### 10. 從矩陣中取數值

先將矩陣 assign 給一個變數，再從中取數值

ex: `mat= matrix(c(1,2,3,4,5,6,7,8,9), nrow=3, byrow=TRUE)`

`mat[1,2]`

>2

因為第一列第二行的數值是 2。

`mat[c(1,3), 2]`

>2 8

因為第一列和第三列的第二行的數值分別是 2 和 8。

`mat[2,]`

>4 5 6

因為第二列的數值分別是 4,5,6。

`mat[,1]`

>1 4 7

因為第一行的數值分別是 1,4,7。

#### 11.矩陣的四則運算

ex: `mat= matrix(c(1,2,3,4,5,6,7,8,9), nrow=3, byrow=TRUE)`

`mat*10`

```
> mat*10
      [,1] [,2] [,3]
[1,]   10   20   30
[2,]   40   50   60
[3,]   70   80   90
>
```

#### 12. 將 Excel 檔放入 R

方法一：

將 Excel 檔存成 csv

```
data1 <- read.csv(file.choose(), header=T)
```

讀取 csv：read.csv

瀏覽檔案並選取：file.choose()

header=T：檔案的第一列有包含變數名稱

header=F：檔案的第一列不包含變數名稱

再輸入 data1

```
> data1
  LungCap Age Height Smoke Gender Caesarean
1   6.475   6  62.1    no   male         no
2  10.125  18  74.7   yes female         no
3   9.550  16  69.7    no female        yes
4  11.125  14  71.0    no   male         no
5   4.800   5  56.9    no   male         no
6   6.225  11  58.7    no female         no
7   4.950   8  63.3    no   male        yes
8   7.325  11  70.4    no   male         no
9   8.875  15  70.5    no   male         no
10  6.800  11  59.2    no   male         no
```

方法二：

將 Excel 檔存成 csv

```
data2 <- read.table(file.choose(), header=T, sep=" ;" )
```

讀取資料：read.table

sep=" ;"：以「;」來做資料分隔依據

再輸入 data2

```
> data2
  LungCap Age Height Smoke Gender Caesarean
1   6.475   6  62.1    no   male         no
2  10.125  18  74.7   yes female         no
3   9.550  16  69.7    no female        yes
4  11.125  14  71.0    no   male         no
5   4.800   5  56.9    no   male         no
6   6.225  11  58.7    no female         no
7   4.950   8  63.3    no   male        yes
8   7.325  11  70.4    no   male         no
9   8.875  15  70.5    no   male         no
10  6.800  11  59.2    no   male         no
> |
```

方法三：

將 Excel 檔存成 txt

```
data3 <- read.delim(file.choose(),header=T)
```

read.delim：讀取 txt 檔

```
> data3
```

	LungCap	Age	Height	Smoke	Gender	Caesarean
1	6.475	6	62.1	no	male	no
2	10.125	18	74.7	yes	female	no
3	9.550	16	69.7	no	female	yes
4	11.125	14	71.0	no	male	no
5	4.800	5	56.9	no	male	no
6	6.225	11	58.7	no	female	no
7	4.950	8	63.3	no	male	yes
8	7.325	11	70.4	no	male	no
9	8.875	15	70.5	no	male	no
10	6.800	11	59.2	no	male	no

方法四：

```
read.table(file.choose(), header=T, sep=" \t" )
```

\t：讓 R 了解這是 txt 檔

再輸入 data4

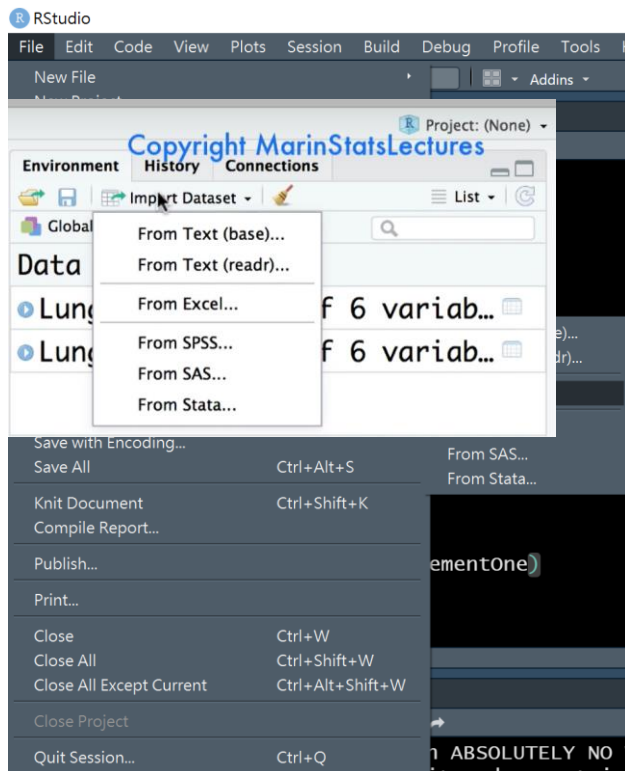
```
> data4
```

	LungCap	Age	Height	Smoke	Gender	Caesarean
1	6.475	6	62.1	no	male	no
2	10.125	18	74.7	yes	female	no
3	9.550	16	69.7	no	female	yes
4	11.125	14	71.0	no	male	no
5	4.800	5	56.9	no	male	no
6	6.225	11	58.7	no	female	no
7	4.950	8	63.3	no	male	yes
8	7.325	11	70.4	no	male	no
9	8.875	15	70.5	no	male	no
10	6.800	11	59.2	no	male	no

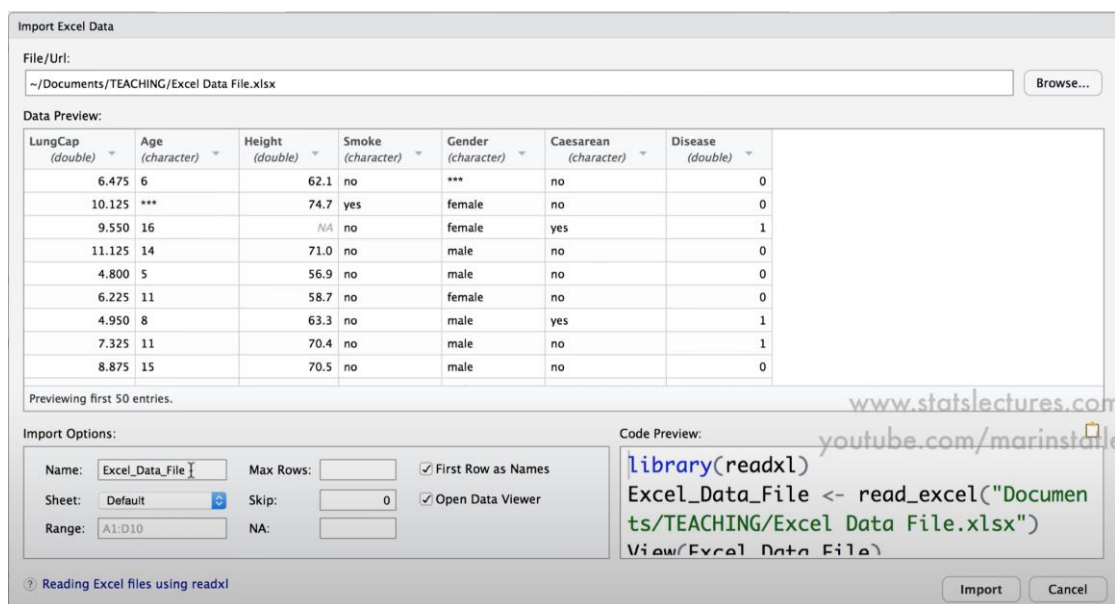
```
>
```

### 13. 輸入部分 excel

可以在 file -> import data -> From Excel 來輸入



或是在 Import Dataset -> From Excel



瀏覽資料夾，找到檔案

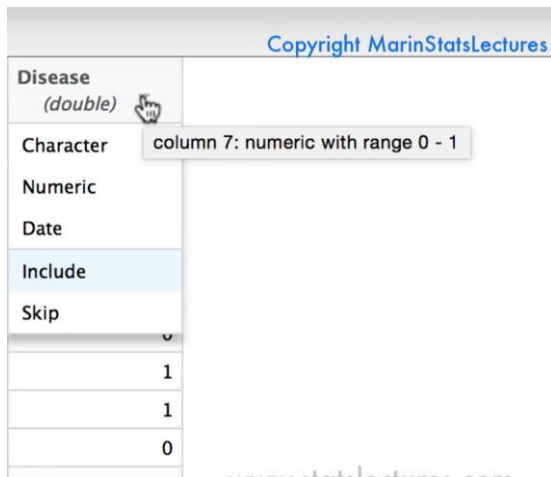
Sheet：選擇試算表 (Default 是第一頁試算表)

Range：資料擷取範圍，輸入左上和右下的儲存格即可，例如 A3:F5

Max Rows：列數上限

Skip：跳掉前幾行

NA：沒有值的代號



可為每行資料辨認屬性或是 Skip。

接著再將右下的程式碼 Import，試算表即出現。

14. 輸出 R 和刪除多餘的編號

```
write.table(DataToExport, file=" ExportedFileName.csv" , sep=" , " )
```

DataToExport：變數名稱（表格的名稱）

ExportedFileName：輸出的檔案名稱

A	B	C	D	E
Subject	Age	Gender	Score	
1	Dave.Andrey	53	male	80.5
2	Jon.Stewart	54	male	82.1
3	Jane.Doe	38	female	75.9
4	Amelia.Earha	119	female	90
5	Donald.Trum	70	male	-25.5
6	Sidney.Crosb	28	male	87.2
7	Oprah.Winfr	62	female	88.8
8	Steve.Jobs	61	male	91.1

打開檔案以後，左邊多了一行數字（12345678），我們可以插入

row.names=F 來去掉它：

```
write.table(DataToExport, file=" ExportedFileName.csv" , row.names=F,
sep=" , " )
```

15. 輸出 R 在指定資料夾

方法一：

```
write.table(DataToExport, file=" /Users/ripple/110-1
/ExportedFileName.csv, row.names=F, sep=" , " )
```

write.table(路徑/檔名, 資料間格依據)

如此即可將檔案輸出成 csv 檔。

方法二：

```
write.csv(DataToExport, file=" /Users/ripple/110-1
/ExportedFileName.csv, row.names=F)
```

方法三：

```
write.table(DataToExport, file="/Users/ripple/110-1/ExportedFileName.txt", row.names=F, sep=" \t" )
```

如此即可將檔案輸出成 txt 檔。

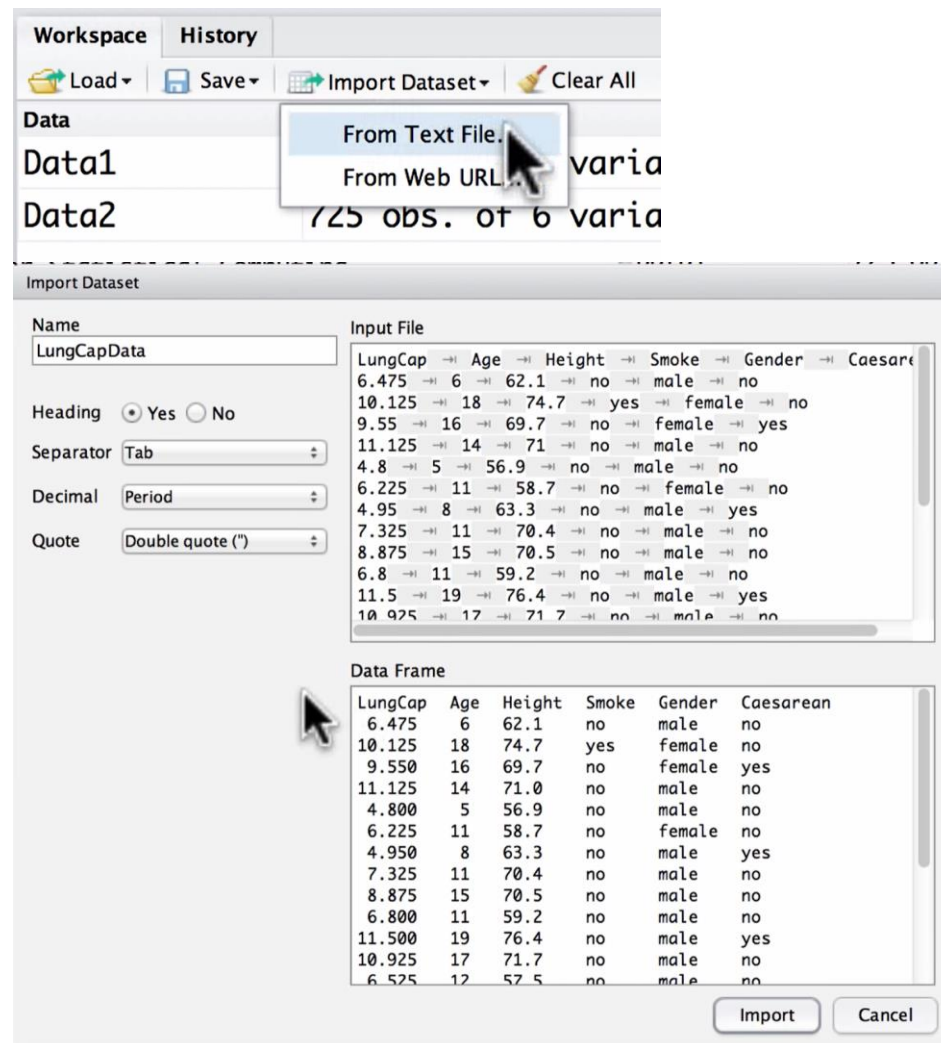
方法四：

```
write.table(DataToExport, file="/Users/ripple/110-1/ExportedFileNameSpace", row.names=F, sep=" ")
```

如此即可將檔案輸出成以空格做資料區隔的檔案。

\*路徑也可以直接用 file.choose()來搜尋

方法五：



Heading：是否包含變數名稱

Separator：資料以什麼作為間隔

Decimal：用什麼來表示小數

Quote：用什麼框起來代表資料類別

接著 Import 就完成。



### 15. 檢視資料數和欄數

`dim(x)`

`> 725 6`

代表有 725 個檢測值、6 個欄位

### 16. 檢視資料頭、尾、或特定列

頭：`head(x)`

尾：`tail(x)`

5~9 列：

`x[c(5, 6, 7, 8, 9), ]`

\*空格代表 all

`x[5:9, ]`

去除 4~722 列：

`x[-(4:722), ]`

### 17. 檢視表格的類別名

`names(x)`

### 18. 顯示表格中類別值並取平均

方法一：

表格名\$類別名

ex: `LungCapData$Age`

`mean(表格名$類別名)`

ex: `mean(LungCapData$Age)`

`>12.3269`

方法二：

`attach(表格名)`：將表格內的所有內容存在 R 的記憶體中

ex: `attach(LungCapData)`

`mean(Age)`

`>12.3269`

`detach(表格名)`：將 R 的記憶體中表格內的內容刪掉

### 19. 檢視 factor 內容

`levels(屬性為 factor 的類別名)`

ex: levels(gender)  
> "female" " male"

## 20. 整理表格內容

summary(表格名)

```
> summary(LungCapData)
```

LungCap	Age	Height	Smoke	Gender
Min. : 0.507	Min. : 3.00	Min. :45.30	no :648	female:358
1st Qu.: 6.150	1st Qu.: 9.00	1st Qu.:59.90	yes: 77	male :367
Median : 8.000	Median :13.00	Median :65.40		
Mean : 7.863	Mean :12.33	Mean :64.84		
3rd Qu.: 9.800	3rd Qu.:15.00	3rd Qu.:70.30		
Max. :14.675	Max. :19.00	Max. :81.80		
Caesarean				
no :561				
yes:164				

\*若屬性為 numeric，則 summary 時會顯示最小值、中位數、平均等等資訊；  
而若屬性為 factor，則 summary 時會顯示兩類型資料的比數。

## 21. 將資料屬性改成 factor

as.factor(變數名)

ex: x <- as.factor(x)

class(x)

> "factor"

## 22. 取出表格中的特定數值和求出其平均

ex: FemData <- LungCapDate[Gender==" female" , ]

\* [] : Subset

計算表格中特定數值的平均

ex: 計算表格中女性的年齡的平均

mean(FemData)

或是

mean(Age[Gender==" female" ])

取出超過 15 歲的女性

ex: FemOver15 <- LungCapDate[Gender==" female" & Age>15, ]

### 23. 增加行或列

`cbind(表格名, Vector)` : 增加一行 Vector

`rbind(表格名, Vector)` : 增加一列 Vector

ex: `FemSmoke <- Gender==" female" & Smoke==" yes"`

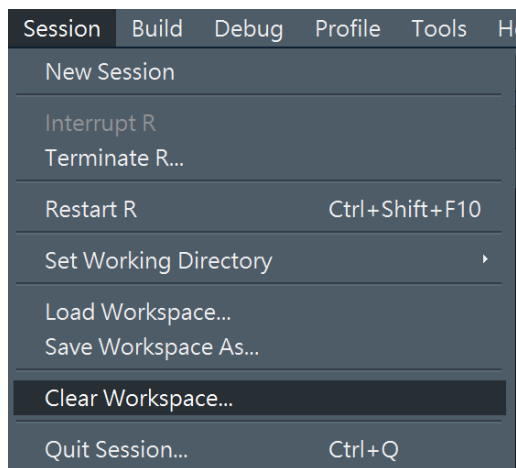
`MoreData <- cbind(LungCapData, FemSmoke)`

```
> MoreData[1:5,]
  LungCap Age Height Smoke Gender Caesarean FemSmoke
1  6.475   6  62.1   no   male         no    FALSE
2 10.125  18  74.7  yes female         no     TRUE
3  9.550  16  69.7   no female        yes    FALSE
4 11.125  14  71.0   no   male         no    FALSE
5  4.800   5  56.9   no   male         no    FALSE
```

### 24. 刪除記憶體中的全部資料

`rm(list=ls())`

(或是 Session -> Clear Workspace )

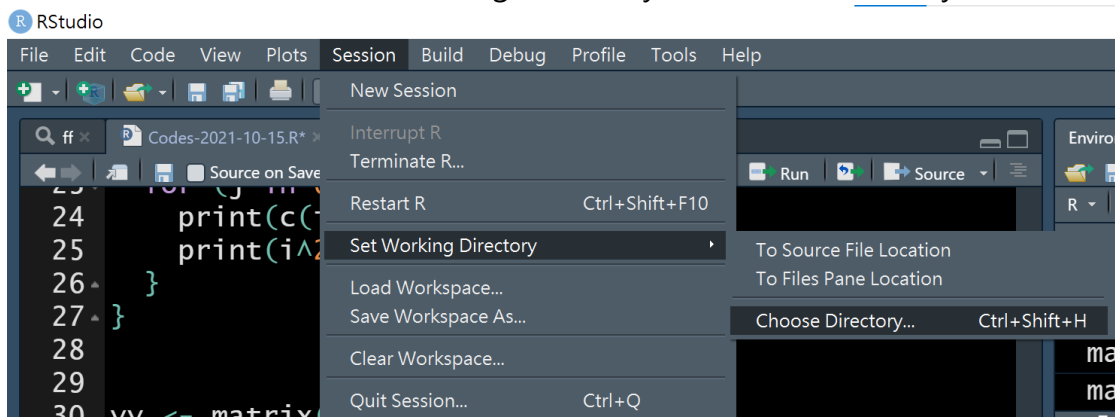


### 25. 找/換工作目錄 working directory

`getwd()` : 找當前的工作目錄

`setwd( "路徑" )` : 切換 working directory 至其他路徑

( 也可以選 Session->Set Working Directory->Choose Directory )



ex: `setwd ( "/Users/OldMarin/Desktop/Project 1" )`

或是 `setwd( "~/Desktop/Project 1" )`

或是 `projectWD <- "/Users/OldMarin/Desktop/Project 1"`

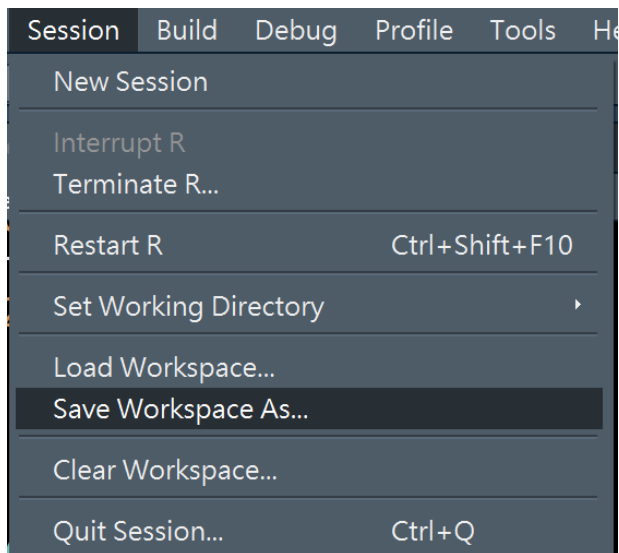
`setwd(projectWD)`

## 26. 儲存 R 檔案、R Script

`save.image( "檔名" )`

\*無指派路徑，則會自動存在工作目錄中。

( 或是選 Session -> Save Workspace As )

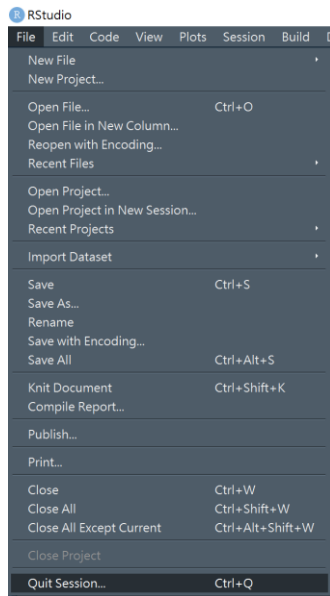


## 27. 關閉 R Studio

`q()`



( 或是 File -> Quit Session )



## 28. 開啟檔案

先設定工作目錄 -> 確認現在在工作目錄 -> 輸入資料

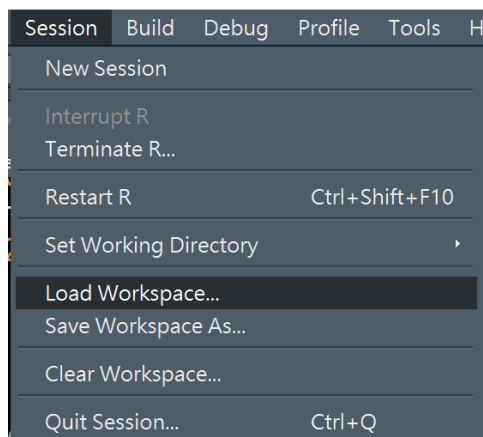
`setwd( "路徑" )`

`getwd()`

`load( "檔名" )`

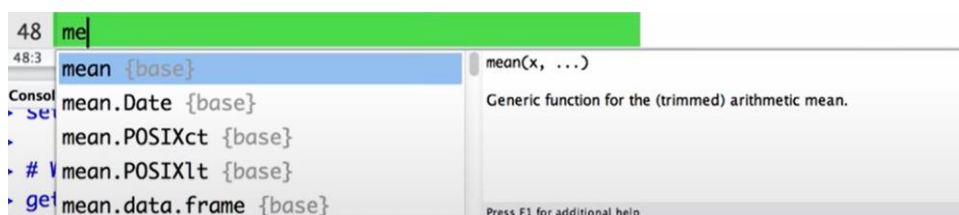
或是 `load(file.choose())` 直接手動找。

( 或是 Session -> Load Workspace )



## 29. 打 code 提示

在 Script 或 Console 中輸入 code 到一半時，按下 `tab` 鍵，就可以找到提示。



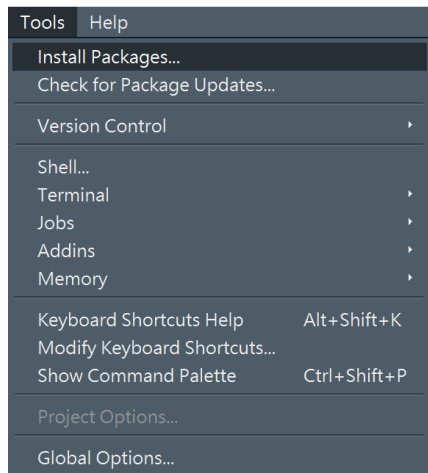
### 30. 下載外掛、使用外掛

`install.packages()`：會顯示所有可下載外掛，可以在自己選擇

`install.packages(“外掛編碼”)`：下載特定外掛

ex: `install.packages(“epiR”)`

( 或是 Tools -> Install Packages )



`library(外掛編碼)`：開啟使用外掛

\* 外掛編碼可以上 <https://cran.r-project.org/>，點左側的 Packages，即可搜尋、挑選。

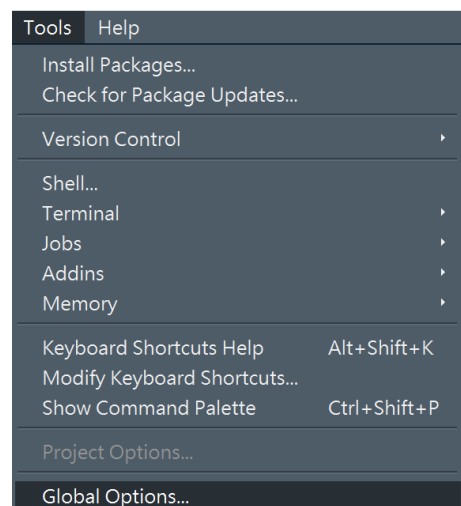
### 31. 外掛說明、移除外掛

`help(packages = 外掛編碼)`

`remove.packages(“外掛編碼”)`

### 32. 改 R Studio 設定

Tools -> Global Options



### 33. apply()

計算表格數字

apply(被指定的表格, margin, 功能, na.rm=TRUE)

margin: 1 代表 row ; 2 代表 column

na.rm=TRUE: 排除 NA 計算平均

apply(StockData, 2, mean) : 按直行計算平均

apply(StockData, 2, max) : 按直行找到最大值

apply(StockData, 2, quantile, probs=c(20,80)) : 按直行找出 20 和 80 分位

apply(StockData, 2, plot, type="l", main=" 標題", ylab=" y 軸名稱",  
xlab=" x 軸名稱" ) : 按直行數值做成折線圖

apply(StockData, 1, sum) : 按橫列計算加總

ex: 表格內容如下 :

	Stock1	Stock2	Stock3	Stock4
Day1	185.74	1.47	1605	95.05
Day2	184.26	1.56	1580	97.49
Day3	162.21	1.39	1490	88.57
Day4	159.04	1.43	1520	85.55
Day5	164.87	1.42	1550	92.04
Day6	162.72	1.36	1525	91.7
Day7	157.89	NA	1495	89.88
Day8	159.49	1.43	1485	93.17
Day9	150.22	1.57	1470	90.12
Day10	151.02	1.54	1510	92.14

求每欄的平均

```
> apply(X=StockData, MARGIN=2, FUN=mean)
      Stock1      Stock2      Stock3      Stock4
163.746      NA 1523.000      91.571
```

\*NA 是因為 Stock2 中有缺空

或是

```
> apply(StockData, 2, mean, na.rm=TRUE)
      Stock1      Stock2      Stock3      Stock4
163.746000      1.463333 1523.000000      91.571000
```

ex: 找出直行中最大值

```
> apply(X=StockData, MARGIN=2, FUN=max, na.rm=TRUE)
```

Stock1	Stock2	Stock3	Stock4
185.74	1.57	1605.00	97.49

```
> apply(X=StockData, MARGIN=2, FUN=quantile, probs=c(0.2, .80),  
+       na.rm=TRUE)
```

	Stock1	Stock2	Stock3	Stock4
20%	156.516	1.408	1489	89.618
80%	168.748	1.548	1556	93.546

34. 計算各項平均 : colMeans()

colMeans(被指定的表格, na.rm=TRUE)

35. 計算各列總和 : rowSums()

rowSums(被指定的表格, na.rm=TRUE)

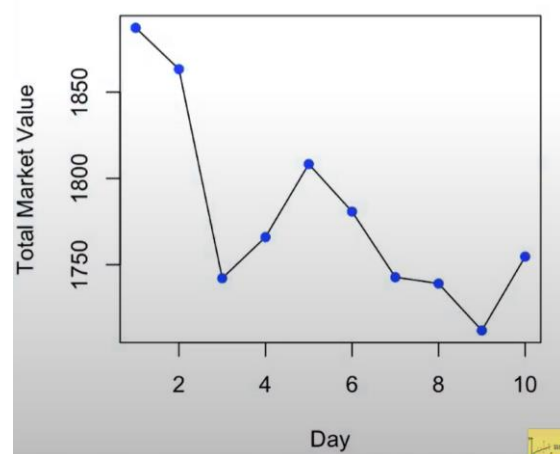
36. 畫圖 : plot() & point()

plot(apply(X, MARGIN, FUN, na.rm=TRUE), type="l", main=" 標題" ,  
ylab=" y 軸名稱" , xlab=" x 軸名稱" ) : 將 apply 中計算出來的表格畫成圖

point(apply(X, MARGIN, FUN, na.rm=TRUE), pch=16, col=blue) : 將 apply  
中計算出來的表格畫成圖並加點

ex:

```
> points(apply(X=StockData, MARGIN=1, FUN=sum, na.rm=TRUE),  
+        pch=16, col="blue")
```



37. tapply() & by()



tapply(X 被函數套用的變數, INDEX 子項目, FUN 功能, na.rm=TRUE 忽略 NA, simplify=TRUE)：對表格中的單一子項目進行另一變數運算，並簡化結果

或是 tapply(X, INDEX=list(子項目 1, 子項目 2), FUN, na.rm=TRUE)

或是 by(X, INDEX=list(子項目 1, 子項目 2), FUN, na.rm=TRUE)

\*只是 by 回傳的形式是向量。

FUN 還可以用 summary、quantil, c(0.2,0.8)

ex: 計算表格中有吸菸和沒有吸菸的人的平均年齡

tapply(X=Age, INDEX=Smoke, FUN=mean, na.rm=T)

```
> tapply(X=Age, INDEX=Smoke, FUN=mean, na.rm=T)
      no      yes
12.03549 14.77922
```

或是 tapply(Age, Smoke, mean)即可

\*以上 yes 的值與 mean(Age[Smoke==" yes" ])一樣)

ex: 未簡化

```
> tapply(Age, Smoke, mean, simplify=FALSE)
$no
[1] 12.03549

$yes
[1] 14.77922
```

ex: 子項目的總括表

```
> tapply(Age, Smoke, summary)
$no
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
 3.00   9.00   12.00   12.04   15.00   19.00

$yes
  Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
10.00  13.00   15.00   14.78   17.00   19.00
```

ex: 子項目的 20、80 分位

```
> tapply(Age, Smoke, quantile, probs=c(0.2, 0.8))
$no
20% 80%
 8  16

$yes
20% 80%
12  17
```

ex: 計算表格中有吸菸和沒有吸菸的男女的平均年齡

tapply(X=Age, INDEX=list(Smoke, Gender), FUN=mean, na.rm=T)

```
> tapply(X=Age, INDEX=list(Smoke, Gender), FUN=mean, na.rm=T)
      female      male
no  12.12739 11.94910
yes 14.75000 14.81818
```